

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-192000

(43)Date of publication of application : 28.07.1995

(51)Int.Cl. G06F 17/28
G06F 17/30

(21)Application number : 05-348536
(22)Date of filing : 27.12.1993

(71)Applicant : TOSHIBA CORP
(72)Inventor : ITO ETSUO
NOMURA KOICHI
TAKEDA KIMITO
MIKE SEIJI

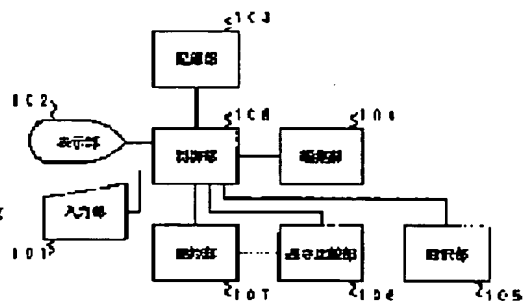
(54) MACHINE TRANSLATION METHOD EQUIPPED WITH SUMMARY FUNCTION

(57)Abstract:

PURPOSE: To easily reflect the layout information of the original sentence to the translated one by comparing the length of the sentence before and after the translation and summarizing the translated sentence as short as the original one when it becomes longer than the original one.

CONSTITUTION: A control part 108 accepts the translation request from an input part 101 such as a keyboard and sends the original sentence to a translation part 105. The part 105 decides the grammatical role of the phrases comprising the input original sentence by performing the morpheme analysis and the syntax analysis and finds out the proper word from the translation dictionary based on the result. Then, and executing a drying operation, based on a result calculated by a time counting means and a variation rate calculating means.

CONSTITUTION: A motor 18 and a heater 13 are connected to a microcomputer 25 through a motor driving circuit 26 and a heater driving circuit 27 integrated into a controller 21, the cloth quantity of an object to be dried is calculated from the time required until an output of a drying degree detecting device 20 becomes a certain prescribed value after the operation is started, and the cloth qger than the original.



LEGAL STATUS

[Date of request for examination] 01.02.2000

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-192000

(43) 公開日 平成7年(1995)7月28日

(51) Int.Cl. ⁸	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 17/28 17/30		8219-5L 9194-5L	G 0 6 F 15/ 38 15/ 401	Z 3 2 0 A

審査請求 未請求 請求項の数 2 F D (全 11 頁)

(21) 出願番号	特願平5-348536	(71) 出願人	000003078 株式会社東芝 神奈川県川崎市幸区堀川町72番地
(22) 出願日	平成5年(1993)12月27日	(72) 発明者	伊藤 悦雄 神奈川県川崎市幸区小向東芝町1番地 株 式会社東芝研究開発センター内
		(72) 発明者	野村 浩一 神奈川県川崎市幸区小向東芝町1番地 株 式会社東芝研究開発センター内
		(72) 発明者	武田 公人 神奈川県川崎市幸区小向東芝町1番地 株 式会社東芝研究開発センター内
		(74) 代理人	弁理士 鈴江 武彦

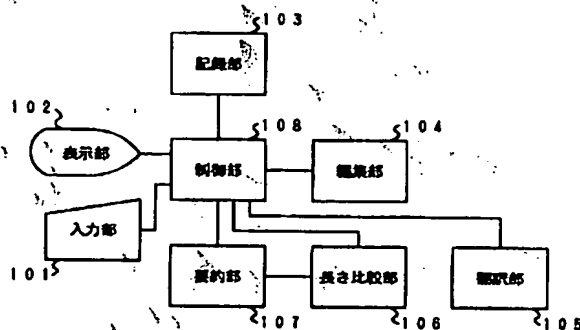
最終頁に続く

(54) 【発明の名称】 要約機能を備えた機械翻訳方法

(57) 【要約】

【目的】 翻訳結果の文章がその原文の文章より長くなった場合に同じ長さになるように要約する機能を有する機械翻訳方法を提供すること。

【構成】 本発明の機械翻訳方法では、第一言語の文章を第二言語の文章に翻訳する翻訳ステップと、翻訳前の第一言語の文章の長さとその翻訳結果である第二言語の文章の長さとを比較する長さ比較ステップと、第一言語の翻訳結果たる第二言語の文章がその原文である第一言語の文章より長くなった場合に、該第二言語の文章を該第一言語の文章とはほぼ同じ長さになるように要約する要約ステップとを有することを特徴としている。



【特許請求の範囲】

【請求項1】第一言語の文章を第二言語の文章に翻訳する翻訳ステップと、

翻訳前の第一言語の文章の長さとその翻訳結果である第二言語の文章の長さとを比較する長さ比較ステップと、第一言語の翻訳結果たる第二言語の文章がその原文である第一言語の文章より長くなった場合に、該第二言語の文章を該第一言語の文章とほぼ同じ長さになるように要約する要約ステップとを有することを特徴とする機械翻訳方法。

【請求項2】第一の言語で記述された文書データベースより文書を検索する検索ステップと、この検索ステップで検索された文章を要約する要約ステップと、

この要約ステップで要約された文書を第二言語に翻訳する翻訳ステップとを有することを特徴とする機械翻訳方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、第一言語を第二言語に効率良く翻訳する機械翻訳装置に関する。

【0002】

【従来の技術】近年、コンピュータシステムを使用して、第一言語の文章を第二言語の文章に自動的に翻訳する機械翻訳装置が注目されている。従来の機械翻訳装置では、形態素解析や構文解析などにより入力原文を構成する語句の文法的な役割を決定し、この結果に基づいて訳語を翻訳辞書から見出し、これを所定の規則にしたがって結合することによって訳文を作成している。

【0003】さらに、そういった翻訳システムで翻訳した訳文に対し、原文のレイアウトを反映し原文のイメージを保持したままの翻訳結果を得る技術も開発されている（特開平4-259057号など）。しかし、翻訳の前後で文章の長さが異なる場合が多く、原文のレイアウトをそのまま反映することができないことがあった。この様な場合、例えば原文ではA4の書式で1ページの資料が訳文では2ページになってしまうなど不具合が発生している。この解決のために、文字の大きさや行間・文字間を変えて無理に元のレイアウトに復元する技術も開発されている（特開昭60-20283号）。しかし、この方法ではかえって文章を読みにくくなる場合が多い。

【0004】また、文章を要約する技術もある（特開平05-186570など）。この技術を用いて翻訳結果を要約し短くすることも可能であるが、従来は翻訳結果を一旦テキストベースに変換し、その後要約システムに入力していたため、操作が繁雑であるだけでなく、レイアウトの復元が困難であったり、適切な要約割合の設定が困難であったりするなどの問題があった。従って、単に従来の機械翻訳装置と文章要約装置を組み合わせ使用するだけでは、満足する結果を得ることはできなかつ

た。

【0005】一方、文書データベースから希望の文書を検索する機能はすでに開発されている。しかし、ユーザーが文書データベースより必要な情報を入手する際、検索結果が母国語（そのユーザーが母国語と同様の容易さで使用できる言語を含む。以下、同様。）でない場合にはそれを読むのに時間がかかる。このため検索後、翻訳してからユーザーに提示する方法も考案されている。しかし、通常データベース検索では本当に必要なものの他に必要でない情報も検索されてしまうので、それらを全部読んで必要であるか否かの判断をするのにも労力を要するなどの問題がある上、本来不要であるデータまですべて翻訳しなければならず、無駄が多かった。

【0006】

【発明が解決しようとする課題】このように、従来の機械翻訳装置における機械翻訳方法では、翻訳の前後で文章の長さが異なることが多く、原文のレイアウトをそのまま反映することができないという問題があった。本発明は、上記課題を解決するためになされたものであり、翻訳結果の文章がその原文の文章より長くなった場合に同じ長さになるように要約する機能を有する機械翻訳方法を提供することを第1の目的とする。

【0007】また、従来の検索機能を備えた機械翻訳装置で使用する機械翻訳方法では、本来不要であるデータまですべて翻訳しなければならず無駄が多いという問題があった。また、ユーザーは検索結果が必要であるか否かの判断をおこなうために文献全体を読まなければならない、効率が悪いと言う問題があった。本発明は、上記課題を解決するためになされたものであり、検索によって得られた文献データのうち要約された部分あるいは構造上重要な部分のみの翻訳結果をユーザーに提示することができる機械翻訳方法を提供することを第2の目的とする。

【0008】

【課題を解決するための手段】上記第1の目的を達成するために、本発明（請求項1）に係る機械翻訳方法では、第一言語の文章を第二言語の文章に翻訳する翻訳ステップと、翻訳前の第一言語の文章の長さとその翻訳結果である第二言語の文章の長さとを比較する長さ比較ステップと、第一言語の翻訳結果たる第二言語の文章がその原文である第一言語の文章より長くなった場合に、該第二言語の文章を該第一言語の文章とほぼ同じ長さになるように要約する要約ステップとを有することを特徴とする。

【0009】一方、上記第2の目的を達成するために、本発明（請求項2）に係る機械翻訳方法では、第一の言語で記述された文書データベースより文書を検索する検索ステップと、この検索ステップで検索された文章を要約する要約ステップと、この要約ステップで要約された文書を第二言語に翻訳する翻訳ステップとを有すること

を特徴とする。

【0010】なお、上記構成の代わりに、第一の言語で記述された文書データベースより文書を検索する検索ステップと、この検索ステップで検索された文章を第二言語に翻訳する翻訳ステップと、この翻訳ステップで翻訳された文書を要約する要約ステップとを有することを特徴としても良い。

【0011】また、上記第2の目的を達成するための他の手段としては、第一の言語で記述された文書データベースより文書を検索する検索ステップと、この検索ステップで検索された文章の論理構造を解析する論理構造解析ステップと、この論理構造解析ステップによる論理構造解析結果に従い、構造上重要である部分のみを第二言語に翻訳するステップとを有することを特徴としても良い。

【0012】

【作用】本発明（請求項1）の機械翻訳方法によれば、翻訳前の第一言語の文章の長さとその翻訳結果である第二言語の文章の長さとを比較し、該第一言語の文章より長くなった場合に、該第二言語の文章を該第一言語の文章とほぼ同じ長さになるように要約するので、翻訳結果を原文とほぼ同じ長さにできる。このため、原文のレイアウト情報などを訳文に容易に反映することができる。

【0013】本発明（請求項2）の機械翻訳方法によれば、文書データベースより検索された第一の言語で記述された文章を、要約してから第二言語に翻訳するので、母国語以外のデータベース検索において、検索結果の要約を母国語で読むことができる。このため、文書検索を効率良く、容易に行うことができる。

【0014】

【実施例】以下、図面を参照しながら本発明の各実施例を説明する。まず、本発明の第1の実施例について説明する。図1は、本実施例に関わる機械翻訳装置の概略的な構成を示すブロック図である。図1に示されるように、この機械翻訳装置は、入力部101、表示部102、記録部103、編集部104、翻訳部105、長さ比較部106、要約部107および制御部108を有する構成となっている。

【0015】入力部101は、文章等のデータや翻訳要求等のコマンドなどを入力するためのものであり、通常キーボードなどが用いられる。表示部102は、入力部101から入力されたデータや翻訳結果などを表示するためのものであり、ビットマップディスプレイなどが用いられる。記録部103は、入力部101から入力されたデータや翻訳結果などを記録するためのものである。

【0016】編集部104は、入力部101から入力されたデータや記録部103に記録されているデータに対する編集処理を行う。翻訳部105は、形態素解析や構文解析などにより入力原文を構成する語句の文法的な役割を決定し、この結果に基づいて訳語を翻訳辞書から見

出し、これを所定の規則にしたがって結合することによって訳文を作成する。

【0017】長さ比較部106は、原文と訳文の長さを比較する。要約部107は、長さ比較部106による原文と訳文の長さ比較結果に従い、原文より訳文の方が長い時には、訳文を原文の長さに要約する。制御部108は、これらの各部の制御やデータの転送を行うためのものである。

【0018】図2は、本実施例に関わる機械翻訳装置の制御部108の動作例を示すフローチャートである。まず、制御部108は、翻訳要求が発生すると原文Aを翻訳部105に転送し翻訳結果Bを得る（ステップS201）。次に、原文Aと翻訳結果Bとを長さ比較部106に転送し長さの比cを得る（ステップS202）。

【0019】この長さの比cが1より大きいかな否か（すなわち訳文の方が原文より長いかな否か）を判定し（ステップS203）、長さの比cが1より大きい場合には、長さの比cと翻訳結果Bを要約部107に転送し、翻訳結果Bを $1/c$ の長さに要約する（ステップS204）。以上の動作により、原文Aと同じ長さの翻訳結果B'を得ることができる。

【0020】ここで、ステップS202における長さの比較は、文章全体を比較対象とすることもできるし、章や節などの論理構造単位、パラグラフやページなどの物理構造単位で比較することもできる。また、表やコラムなど、文字の記入できる最大長が規定される場合には、その最大長と比較することもできる。長さの単位は文章データの文字数の他、バイト数、その言語の文字の平均横幅と文字数の積などを用いても良い。例えば、日本語を英語に翻訳した場合、英語の文字（アルファベット）は日本語の文字（かな、漢字）の半分の幅で表現されることが多いため、文字数が2倍以上になった場合に始めて英文の方が長いとすることもできる。

【0021】ここで、文の長さの調整処理では、フルセルで記述されているものを略号に変更する方法を含んでも良い。この方法について、「機械翻訳」という語句を日英翻訳する場合を例にして説明する。図3は、日英翻訳における原文中に「機械翻訳」という語句が出現し、「機械翻訳」の約語として、「MT」、「machine translation」、「natural language translation system」がある場合の例を示している。図4は、訳語の長さの調整を行うためのデータベースの例を図3に示す例で示したものである。このデータベースでは、それぞれの訳語に、訳語を生成する文字列の長さを同時に記録している。図5には、図3および図4に示す例を用いて、訳文の長さを調整する翻訳システムにおいて、訳語の選択を行う動作の例を示す。なお、このフローチャート中、パラメータiは、長さ調整する文章中に並ぶ語句のうち、処理対象となる語句の先頭からの順序を表す。

まず、訳語の選択は原文の各単語に対して行うため、原文の先頭の単語に着目する(ステップS301)。その単語に対応する訳語が1つしか無い場合は訳語の選択の余地がないため、次の原文の単語に注目語を移す(ステップS303~S304)。複数の訳語が存在する場合には(ステップS303)、各訳語の長さを調整する。このため、訳語長さを調整しないときに出力される翻訳結果では希望する出力文の長さより短くなる場合には、長さを調整しないときに出力される訳語より長い訳語を選択し、一方長くなる場合には、短い単語を選択する(ステップS305)。そして、処理対象の語句を、この選択した語句に置き換える(ステップS306)。この選択により出力される文の長さが希望する長さと比較し許容範囲に含まれる場合には(ステップS307)、もはや長さを調整する必要がないので、処理を終える。長さの差が許容範囲を越える場合には(ステップS307)、さらに調整を行うために、次の単語に注目する語を移動し(ステップS304)、これを文の最後の単語まで繰り返す。

【0022】このように、翻訳を行った結果に対して長さの調整を行うと、原文と調整前の翻訳結果のずれ(長さの差)が容易に分かるため、どの程度の長さ調整(要約)をおこなえば良いのかの判断がしやすく、効率的な長さ調整を行うことができる。

【0023】さらに、長さの比較は、文章中にキーワードが出現する位置を基準に行い、文頭からその位置までの長さ同士を比較しても良い。例えば、図面を参照する文字列(“fig. 1.”、“第1図”など)の位置が翻訳前後で同じになるように要約する長さを調整することにより、参照箇所と被参照箇所の位置関係が原文と訳文で同じになり、読み易い翻訳結果を得ることができる。この例を図6、図7および図8に示す。図6に示すような翻訳の原文となる文書を、従来の翻訳システムを用いて翻訳し、レイアウト復元を行うと図7のように2ページとなってしまい、さらに図面の参照箇所が2ページ目になってしまっている。これに対して、本実施例に関わる機械翻訳装置で翻訳を行うと図8のように原文とほぼ同様のレイアウトとなり、読みやすい翻訳結果が得られる。

【0024】次に、本発明の第2の実施例について説明する。図9は、本実施例に関わる機械翻訳装置の概略的な構成を示すブロック図である。図9に示されるように、この機械翻訳装置は、入力部601、表示部602、記録部603、編集部604、検索部605、文書データベース606、要約部607、翻訳部608および制御部609を有する構成となっている。

【0025】入力部601は、データや検索開始等のコマンドなどを入力するためのものであり、通常キーボードなどが用いられる。表示部602は、入力部601から入力されたデータや翻訳結果などを表示するためのもの

でありビットマップディスプレイなどが用いられる。記録部603は、入力部601から入力されたデータや翻訳結果などを記録するためのものである。

【0026】編集部604は、入力部601から入力されたデータや記録部に記録されているデータに対する編集処理を行う。検索部605は、指示された条件に合致するデータを文書データベース606より検索する。要約部607は、検索部605で検索されたデータを要約する。

【0027】翻訳部608は、要約部607で要約されたデータを翻訳したり、必要な場合には検索部605で検索されたデータをすべて翻訳する。制御部609は、これらの各部の制御やデータの転送を行う。図10は、本実施例に関わる機械翻訳装置の制御部609の動作を示すフローチャートである。

【0028】まず、制御部609は、入力部601より検索開始コマンドが入力されると動作を開始し、検索条件Aを検索部605へ転送し検索結果Bを得る(ステップS701)。この検索結果Bが複数文書であった場合には、その中から必要な文書を選択するために各文書の記述内容を確認する必要があるため、オペレータにどの文書から中身の確認を行うかの選択をさせ、その文書をCとする(ステップS702、703)。一方、検索結果Bが単一文書である場合には、その記述内容から検索結果Bが希望の文書であるか否かを確認する必要があるため、これをCとする(ステップS704)。

【0029】次に、文書Cを要約部607へ転送し、要約結果Dを得る。要約結果Dがあらかじめ登録されていたオペレータの母国語で記述されているか否かを確認し(ステップS706)、母国語でなければ要約結果Dを翻訳部608へ転送し、翻訳結果Eを得る(ステップS707)。母国語の場合は要約結果DをそのままEとする(ステップS708)。そして、Eをオペレータに提示して動作を終了する(ステップS709)。

【0030】ここで、図11に、本実施例における画面表示の一例を示す。図11において、最上部は検索条件入力領域であり、ここでは「computer」というキーワードでの検索要求がなされている。図11の中段は上記検索条件で検索された文書のタイトルのリストが表示されており、no. 2の文書の内容確認が指示されている。また、下部には内容が提示されているが、要約後翻訳された結果が表示されているので、ユーザは母国語である日本語でこの文書の要旨だけを読むことができる。

【0031】なお、上記構成の代わりに、翻訳後に要約を行う構成を採用することも可能である。ただし、要約後翻訳を行う構成を採用する場合は、要旨以外の部分を翻訳する手間が省かれているため処理時間を短くすることができる、という利点がある。次に、本発明の第3の実施例について説明する。

【0032】図12は、本実施例に関わる機械翻訳装置の概略的な構成を示すブロック図である。図12に示されるように、この機械翻訳装置は、入力部901、表示部902、記録部903、編集部904、検索部905、文書データベース906、文書構造解析部907、翻訳部908および制御部909を有する構成となっている。

【0033】入力部901は、データや検索開始等のコマンドなどを入力するためのものであり、通常キーボードなどが用いられる。表示部902は、入力部901から入力されたデータや翻訳結果などを表示するためのものであり、ビットマップディスプレイなどが用いられる。記録部903は、入力部901から入力されたデータや翻訳結果などを記録するためのものである。

【0034】編集部904は、入力部901から入力されたデータや記録部903に記録されているデータに対する編集処理を行う。検索部905は、指示された条件に合致するデータを文書データベース906より検索する。

【0035】文書構造解析部907は、検索部905で検索された文書データの構造解析をするためのものであり、例えば文献“A Discourse Structure Analyser for Japanese Text” (K. Sumita, Int. conf. on Fifth Generation computer '92)などに記載された手法により文書の構造を得る。

【0036】翻訳部908は、文書構造解析部907で解析され、重要であると判断された箇所のデータを翻訳したり、必要な場合には検索部905で検索されたデータをすべて翻訳する。制御部909は、これらの各部の制御やデータの転送を行う。図13は、本実施例に関わる機械翻訳装置の制御部909の動作を示すフローチャートである。

【0037】まず、制御部909は、入力部901より検索開始コマンドが入力されると動作を開始し、検索条件Aを検索部905へ転送し検索結果Bを得る(ステップS1001)。この検索結果Bが複数文書であった場合には、その中から必要な文書を選択するために各文書の記述内容を確認する必要があるため、オペレータにどの文書から中身の確認を行うかの選択をさせ、その文書をCとする(ステップS1002、1003)。一方、検索結果Bが単一文書である場合には、その記述内容から検索結果Bが希望の文書であるか否かを確認する必要があるため、これをCとする(ステップS1004)。

【0038】次に、文書Cを文書構造解析部907へ転送し、構造解析結果Dを得る。そして、構造解析結果Dより構造上重要であると判断された部分を抽出しこれをEとする(ステップS1006)。Eがあらかじめ登録されていたオペレータの母国語で記述されているか否か

を確認し(ステップS1007)、母国語でなければEを翻訳部へ転送し、翻訳結果Fを得る(ステップS1008)。母国語の場合はEをそのままFとする(ステップS1009)。そして、Fをオペレータに提示して動作を終了する(ステップS1010)。

【0039】ここで、図14に、文書構造解析部907による構造解析の結果の一例を示す。図14において“<”および“>”で囲まれた部分が解析された結果の役割を示す部分であり、それに続いてその役割を果たす文章が続く。ステップS1006における構造上重要であるか否かの判断は、例えば、文章の結論部分であるか、タイトルであるかなどで行い、例えば、結論部分だけを抽出し翻訳してオペレータに提示することができる。この判断基準はオペレータの指示で変更してもよい。

【0040】なお、上記実施例は日本語を母国語として説明しているが、それ以外の言語でも良い。ようするに、本発明はその主旨を逸脱しない範囲において種々変更して実施することができる。

【0041】

【発明の効果】本発明(請求項1)の機械翻訳方法によれば、翻訳結果と原文が同じ長さになるため、原文のレイアウト情報などを訳文に容易に反映することができる。また、本発明(請求項2)の機械翻訳方法によれば、母国語以外のデータベース検索において、検索結果の要約を母国語で読むことができるため、文書検索を効率良く、容易に行うことができる。

【図面の簡単な説明】

【図1】本発明の第1の実施例に関わる要約機能付き機械翻訳装置の概略構成を示すブロック図

【図2】図1の機械翻訳装置の動作の流れを示すフローチャート

【図3】日英翻訳において原文の単語に対応する訳語が複数ある例を示す図

【図4】訳語の長さの調整を行うためのデータベースの一例を示す図

【図5】訳語の選択を行う動作の一例を示すフローチャート

【図6】同実施例を適用する原文の一例を示す図

【図7】従来の機械翻訳方法による翻訳結果の一例を示す図

【図8】同実施例による翻訳結果の一例を示す図

【図9】本発明の第2の実施例に関わる検索・要約機能付き機械翻訳装置の概略構成を示すブロック図

【図10】図9の機械翻訳装置の動作の流れを示すフローチャート

【図11】同実施例による画面表示例を示す図

【図12】本発明の第2の実施例に関わる検索・構造解析機能付き機械翻訳装置の概略構成を示すブロック図

【図13】図12の機械翻訳装置の動作の流れを示すフ

ローチャート

【図14】同実施例による構造解析結果の一例を示す図

【符号の説明】

101…入力部

表示部

103…記録部

編集部

105…翻訳部

長さ比較部

107…要約部

制御部

601…入力部

表示部

603…記録部

編集部

102…

104…

106…

108…

602…

604…

*

*605…検索部

文書データベース

607…要約部

翻訳部

609…制御部

入力部

902…表示部

記録部

904…編集部

108…

906…文書データベース

文書構造解析部

908…翻訳部

制御部

606…

608…

901…

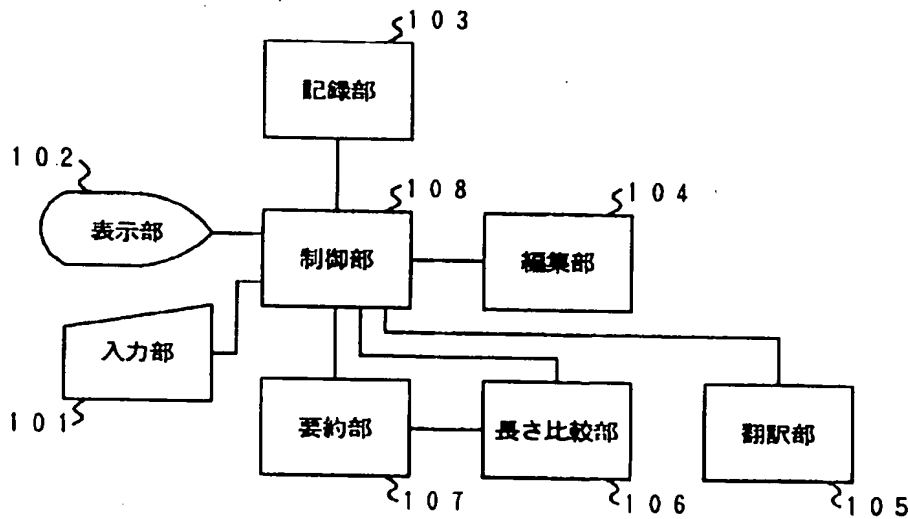
903…

905…

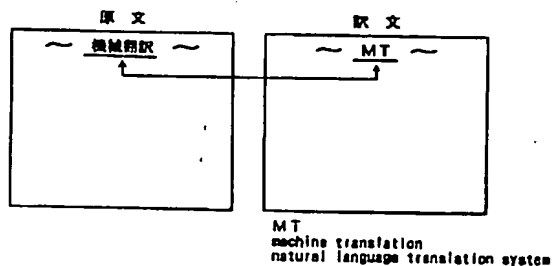
907…

909…

【図1】



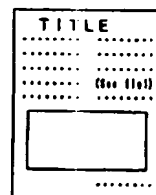
【図3】



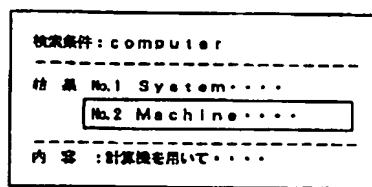
【図4】

〈形式〉 原語 : 訳語の並び
 〈例〉 機械翻訳: (2)MT, (19)machine_translation,
 (34)natural_language_translation_system

【図8】

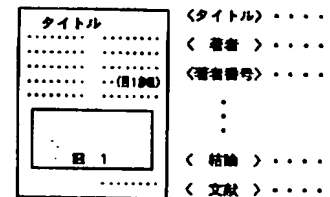


【図11】

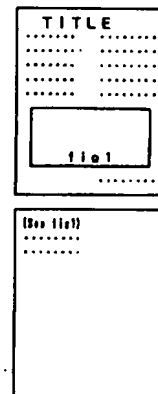


【図6】

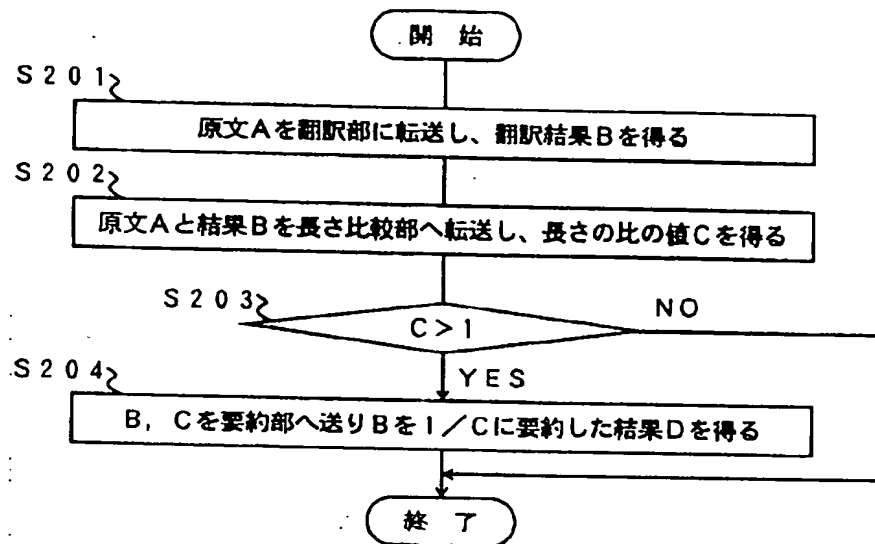
【図14】



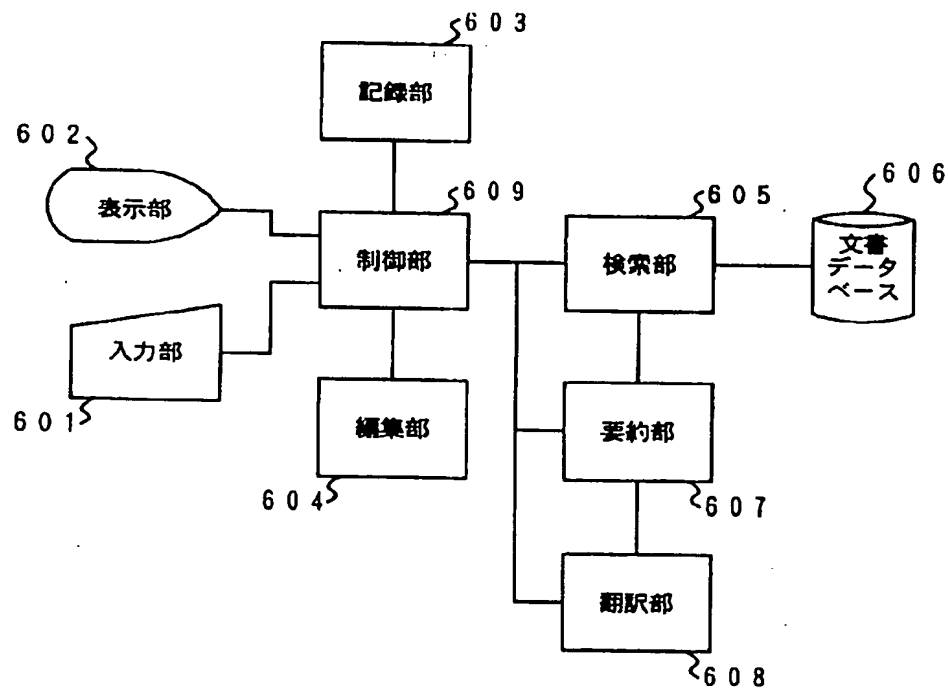
【図7】



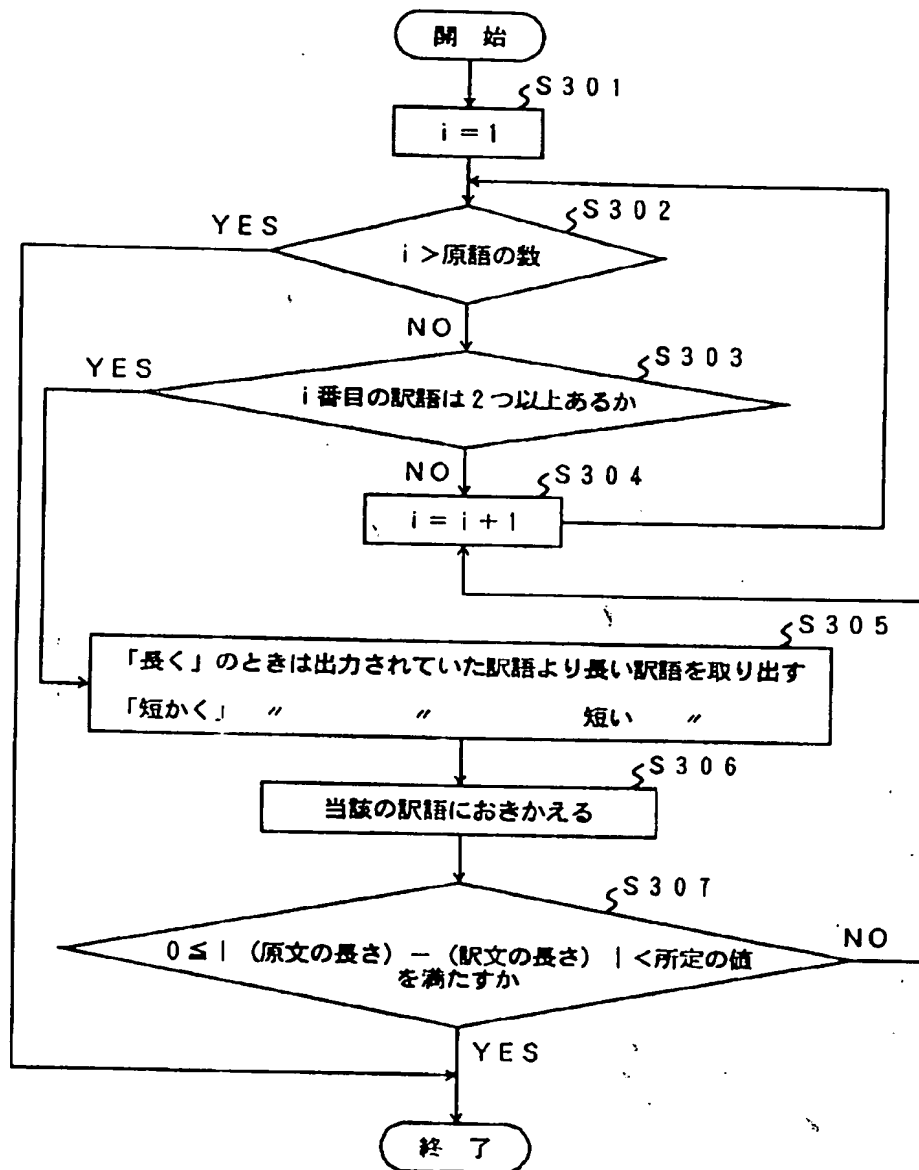
【図2】



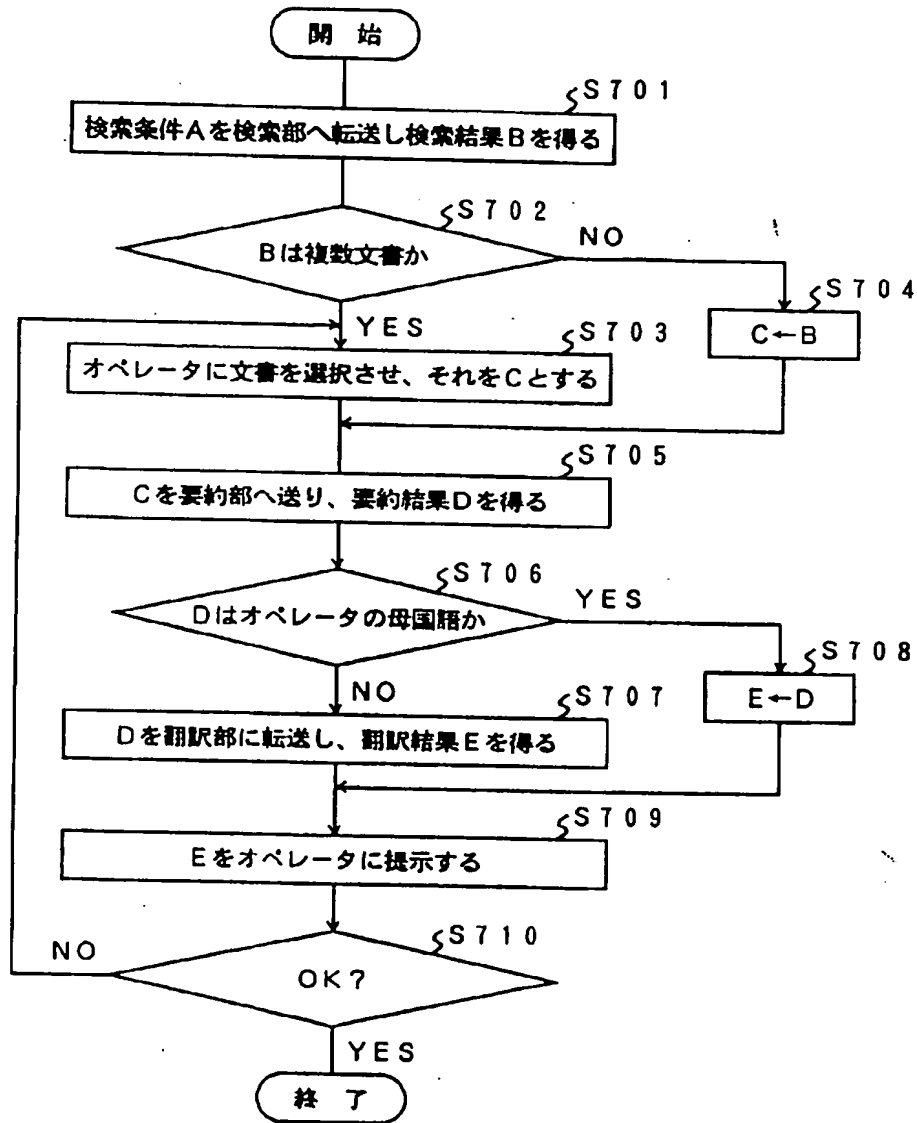
【図9】



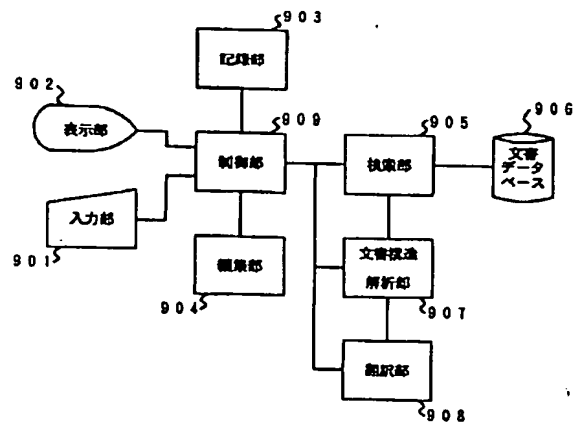
【図5】



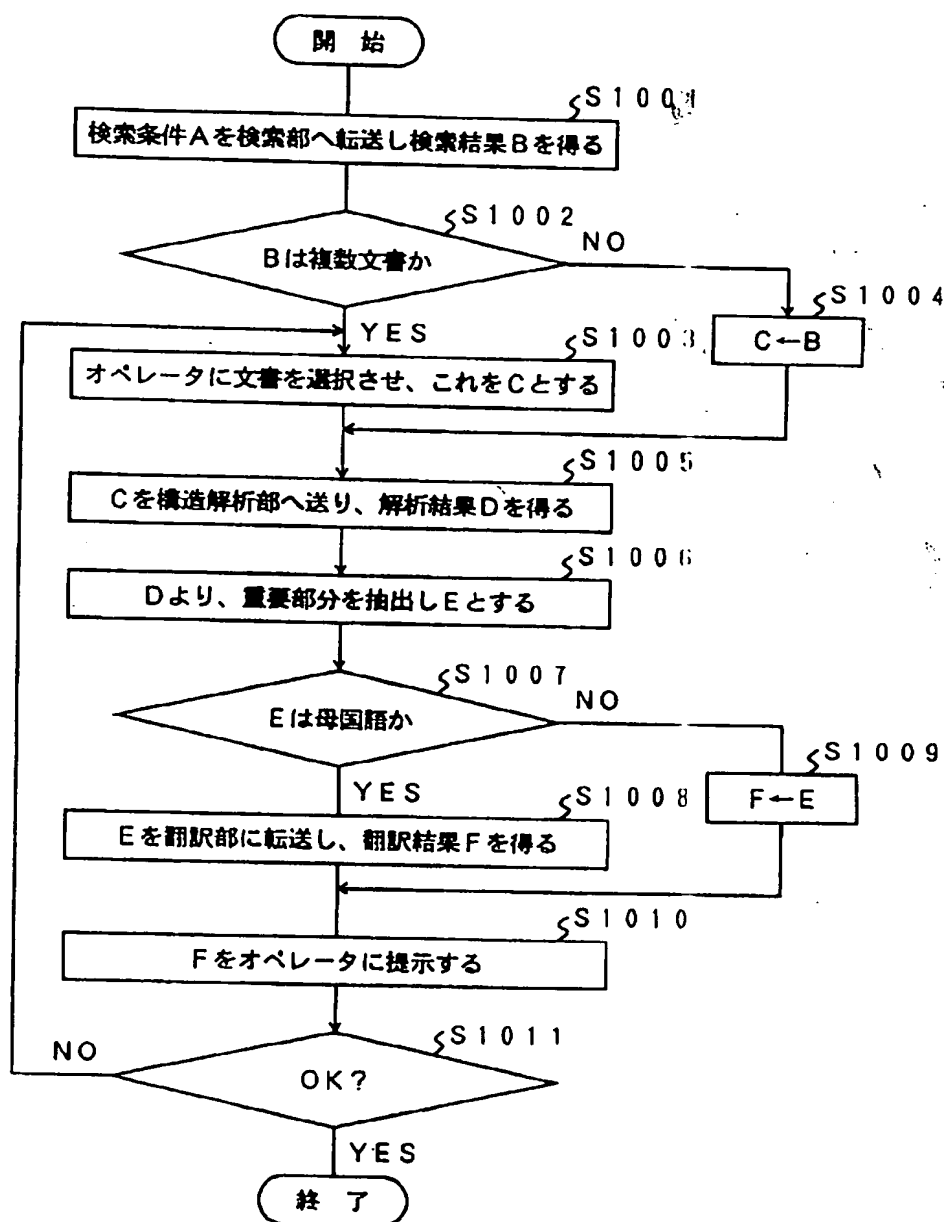
【図10】



【図12】



【図13】



フロントページの続き

(72)発明者 三池 誠司

神奈川県川崎市幸区小向東芝町1番地 株
式会社東芝研究開発センター内